(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 033 374 A1

(12) **EUROPEAN PATENT APPLICATION**
published in accordance with Art. 158(3) EPC

(43) Date of publication:
06.09.2000 Bulletin 2000/36

(21) Application number: **98955942.2**

(22) Date of filing: **26.11.1998**

(51) Int. Cl.[7]: **C07K 14/715**, C12N 15/12, C12N 5/16, C12Q 1/68

(86) International application number:
PCT/JP98/05326

(87) International publication number:
WO 99/26978 (03.06.1999 Gazette 1999/22)

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE

(30) Priority: **26.11.1997 JP 32491297**

(71) Applicants:
- **CHUGAI SEIYAKU KABUSHIKI KAISHA**
  Tokyo, 115-8543 (JP)
- **Kitamura, Toshio**
  Minato-ku, Tokyo 108-0072 (JP)

(72) Inventors:
- **KITAMURA, Toshio**
  Tokyo 108-0072 (JP)
- **KOJIMA, Tetsuo,**
  Chugai Res. Inst. Molec. Med. Inc
  Niihari-gun, Ibaraki 300-4101 (JP)

(74) Representative:
VOSSIUS & PARTNER
Siebertstrasse 4
81675 München (DE)

(54) SIGNAL SEQUENCE TRAPPING METHOD

(57) A DNA encoding human mpl lacking the secretory ability, prepared by deleting the secretory signal of homeostatically active mpl and the major part of the extracellular region thereof. This DNA is ligated with a test cDNA, a cDNA encoding a known secretory protein and another cDNA encoding the same protein from which the secretory signal region is removed, each of the resulting chimeric genes is expressed in cells to examine cell proliferation capability. Cell proliferation is detected in the cells containing the cDNA encoding the known secretory protein as a test cDNA, whereas no cell proliferation is detected for the cDNA encoding the same protein from which the secretory signal region is removed. Thus, cDNAs encoding secretory proteins including type I membrane proteins and type II membrane proteins can be detected and isolated by constructing a cDNA library and screening it by the above method.

EP 1 033 374 A1

Description

Technical Field

[0001] The present invention belongs to the field of gene engineering and relates to a method for detecting and isolating a DNA encoding a peptide that has secretory ability.

Background Art

[0002] So far, genes encoding a hormone or growth factor have been isolated and utilized to produce many recombinant proteins that are commercialized as medicines. Most of them are secretory proteins. Therefore, isolation of a gene encoding a novel secretory protein is an extremely important step in developing a novel medicine. Accordingly, methods for isolating a gene encoding a secretory protein have been developed. For example, Honjo et al. developed a method (unexamined published Japanese patent application No. Hei 6-315380) by utilizing the feature that secretory proteins have a signal sequence that allows intracellularly expressed proteins to translocate to the cell surface. In this method, the signal sequence of the α chain of human IL-2 receptor, a secretory protein, is replaced with a short cDNA fragment corresponding to the 5'-end sequence of mRNA from a target cell or tissue to construct a library, which is then introduced into cells. Among the clones, IL-2 receptor is expressed on the cell surface of clones with a signal sequence, but not those without a signal sequence. The presence of the signal sequence can thus be detected by the anti-IL-2 receptor antibody.

[0003] Genetics Institute, Inc., (Cambridge, MA) developed a more sophisticated system utilizing a yeast metabolic enzyme (U. S. Patent 5,536,637). Invertase, a yeast metabolic enzyme, is a secretory enzyme that cleaves sucrose in the culture medium to glucose and fructose to transfer energy. A mutant strain that does not secrete this enzyme can not grow in a medium containing sucrose as the sole carbon source without glucose. In this method, which utilizes the phenomenon, invertase gene is ligated with cDNA to construct a library, which is then introduced into a mutant yeast strain lacking the invertase gene. Clones containing the signal peptide are isolated by selecting clones capable of growing in a medium containing only sucrose as a carbon source.

[0004] However, the method of Honjo et al. is disadvantageous in that laborious steps are required in selecting positive clones due to the use of an antibody. Furthermore, the detection sensitivity is very low. The method of Genetics Institute, Inc., also has a problem in that a clone with poor secretion efficiency in yeast cannot be isolated. In addition, these methods detect only short DNA because of the potential loss of antigenicity or enzymatic activity when the reporter protein is fused with a large protein. Moreover, the methods fail to detect the type II membrane proteins having their N-terminus within the cell and C-terminus outside the cell.

Disclosure of the Invention

[0005] The present invention provides a method for examining whether a tested cDNA encodes a peptide with the secretory ability or not. It also provides a method for isolating a cDNA encoding a peptide with the secretory activity, which permits the use of a cDNA encoding a long peptide coding region.

[0006] Proteins such as cytokine receptors translocate to the cell surface, dimerize upon binding their ligands and induce cell proliferation. The translocation ability (secretory ability) to the cell surface of the proteins is known to depend on the presence of the signal sequence. The present inventors thought it possible to examine whether a desired peptide has the secretory activity by removing the signal sequence (or with additional extracellular region) from the proteins and replacing it with a fusion protein containing a desired peptide, expressing the fusion protein in cells, and examining the proliferation ability of the cells. If the peptide has the secretory ability, the fusion protein translocates to the cell surface, dimerizes, and induces cell proliferation. In contrast, if the peptide does not contain the secretory activity, the fusion protein cannot translocate to the cell surface and induce cell proliferation. Thus, the secretory ability of the fused peptide can be tested by simply examining cell proliferation as an index. Moreover, the inventors thought it possible to perform positive screening for a peptide with the secretory ability by selecting cells that proliferate. Thus, the present inventors used mpl (thrombopoietin receptor) as a protein that triggers cell proliferation through translocation to the cell surface and dimerization, and developed a method for detecting and isolating a peptide possessing the secretory ability.

[0007] Specifically, we prepared a DNA encoding human mpl without the secretory ability by removing the secretion signal and most of the extracellular domain from a constitutively active form of mpl, which was found by the present inventors (the mpl is altered to be able to confer autonomous proliferation ability to an IL-3 dependent cell line by the transducing signal in the absence of ligand; Blood 88:1399-1406 (1996)). The DNA was then ligated with cDNA to be tested, a DNA encoding a known secretory protein, or a DNA encoding the secretory peptide from which the secretory signal region was removed. The resulting chimeric genes were expressed in cells, and the proliferation ability of the cells was examined. The results show that the DNA encoding a known secretory protein used as a test cDNA induced

cell proliferation whereas no cell proliferation was detected for the DNA encoding the secretory protein from which the secretory signal region was removed. In this way, the inventors found that the system thus developed can be used to easily detect and isolate a DNA encoding a peptide with secretory activity and containing a long peptide coding region using cell proliferation as an index. Indeed, they performed a screening and succeeded in detecting and isolating DNAs encoding secretory proteins including type I membrane proteins and type II membrane proteins.

[0008] Thus, the present invention relates to:

(1) a peptide capable of inducing cell proliferation through dimerization on the cell surface and lacking the secretory ability;
(2) the peptide as described in (1), wherein the peptide is derived from a cytokine receptor;
(3) the peptide as described in (1), wherein the peptide is derived from mpl;
(4) the peptide as described in (2) or (3), wherein the peptide is ligand-independent;
(5) the peptide as described in (1), wherein the peptide comprises the amino acid sequence of SEQ ID NO: 4;
(6) a DNA encoding the peptide as described in any of (1) to (5);
(7) a vector containing the DNA as described in (6) and a cloning site for cDNA at the 5'-upstream region of the DNA;
(8) the vector as described in (7), wherein the vector is derived from a retrovirus;
(9) the vector described in (7) or (8), wherein a cDNA is inserted into the 5'-upstream of the DNA of (6);
(10) a cell carrying the vector as described in (9);
(11) a cell as described in (10), wherein the cell is a mammalian cell;
(12) a method for examining whether a peptide encoded by a cDNA to be tested contains the secretory ability, the method comprising

(a) ligating the test cDNA with the vector of (7),
(b) introducing the vector prepared in (a) into a cell, and
(c) culturing the transformant prepared in (b), and detecting the cell proliferation ability;

(13) a method for isolating a cDNA encoding a peptide with the secretory ability, the method comprising

(a) ligating a cDNA library with the vector of (7),
(b) introducing the vector prepared in (a) into a cell,
(c) culturing the transformant prepared in (b), and detecting the cell proliferation ability, and
(d) selecting a positive cell that is judged to have cell proliferation ability in (c), and isolating the cDNA from said cell;

(14) a method as described in (12) or (13), wherein the vector is derived from a retrovirus and the cell to be introduced with the vector is a mammalian cell;
(15) a cDNA encoding a peptide with the secretory ability, which is isolated by the method of (13); and
(16) a peptide encoded by the cDNA as described in (15).

[0009] The present invention relates to a method for detecting a DNA encoding a peptide having the secretory ability. The detection method features the use of a DNA encoding a peptide capable of inducing cell proliferation through dimerization on the cell surface and lacking the secretory ability for detecting a peptide with the secretory ability. Here, "the peptide capable of inducing cell proliferation through dimerization on the cell surface" includes mpl (Proc. Natl. Acad. Sci. USA, 89:5640-5644 (1992)), the alpha chain or beta chain of GM-CSF (Blood 83:2802 (1994)), erythropoietin receptor (Nature 348:647 (1990)), c-kit receptor (Blood 85:790 (1995)), and neu (Nature 339:230 (1989)), but are not limited thereto. In the method of the present invention, a cDNA is constructed to encode the above peptides whose secretory ability is eliminated. The secretory ability is usually removed by deleting a region containing the signal sequence. For example, the signal sequence of the human mpl is the region corresponding to 1 to 25 positions in the amino acid sequence of the protein (Proc. Natl. Acad. Sci. USA, 89:5640-5644 (1992)), and that of the beta chain of the human GM-CSF is the region corresponding to 1 to 48 positions (Proc. Natl. Acad. Sci. USA, 87:9655-9659 (1990)). Preferably, the extracellular domain is also deleted from the peptide.

[0010] The peptide encoded by a constructed cDNA is preferably ligand-independent (if the peptide is ligand-dependent, it may lose the ligand-binding ability and become inactive after creating a fusion protein). A method for creating a ligand-independent peptide is to introduce a mutation into the peptide, for example. In case of mpl, the substitution of Ser 498 to Asn can abolish the dependency to the ligand, thrombopoietin (Blood 88:1399-1406 (1996)). The mpl used in the present method preferably comprises the amino acid sequence as described in SEQ ID NO: 4.

[0011] The DNA prepared as described above is inserted into an appropriate expression vector. The expression

vector is not limited, and preferably is a retrovirus vector, which can be introduced into a variety of cells with high efficiency through virus infection, and stably expresses the DNA inserted into the vector in the cells. Examples of a retrovirus vector include that engineered for cDNA library construction, such as pBabeX (Proc. Natl. Acad. Sci. USA 92:9146-9150 (1995)) or pMX (Exp. Hematol. 24:324-329 (1996)). Also, virus vectors such as adenovirus, EB virus, and papilloma virus, or plasmid vectors such as pEF-BOS (Nucleic Acid Res. 18 (17)) and pcD SRα296 (Mol. Cell. Biol. Jan. 466-472 (1988)) can be used. The expression vectors should have a cloning site for a cDNA to be tested for its secretory ability at the 5'-upstream of the above DNA insert to express a fusion protein. The method for creating a cloning site for a cDNA is known to one skilled in the art.

[0012] Next, the prepared vector is ligated with a cDNA to be tested. The test cDNA is ligated into the 5'-upstream of the "DNA encoding a peptide capable of inducing cell proliferation through dimerization on the cell surface," which is inserted into the vector. The test cDNA can be any cDNA encoding a peptide whose secretory ability is to be tested. The test cDNA can be ligated with a vector according to the standard methods. For example, the ligation method using T4 DNA ligase via an adapter • linker (Maniatis T., Molecular Cloning).

[0013] The prepared vector is then introduced into a cell. Cells into which the vector is introduced are not limited and include cytokine-dependent proliferating cell such as Ba/F3, OTT-1, FDCP-1, and TF-1 cells. Vectors can be introduced into cells by using standard methods, including lipofection, calcium phosphate method, DEAE-dextran method, and electroporation. In retrovirus infection-mediated introduction, the vector is introduced into the packaging cells and integrated into the virus particles. The vector can be introduced by using standard methods such as the calcium phosphate method and lipofection. For example, cells such as BOSC23, Bing (Proc. Natl. Acad. Sci. USA 90:8392-8396 (1993)), NX-E, and NX-A cells (Nolan G.P. Immunity 8:461-471 (1998)) can be used as the packaging cell.

[0014] Next, the thus-prepared transformants are cultured and examined for their proliferation ability. When a protein encoded by cDNA inserted into the vector is expressed as a fusion protein with a ligand-independent active cytokine receptor, the transformant is cultured in the medium lacking the cytokine (ligand) on which the cell depends. If a significant cell proliferation is detected, the test cDNA is judged to be a "positive clone" encoding a peptide that contains the secretory ability. Alternatively, if no significant cell proliferation is detected, the cDNA is judged to be a "negative clone" that encodes a peptide lacking the secretory ability. When a protein encoded by the inserted cDNA is expressed as a fusion protein with a ligand-dependent cytokine receptor, the transformant is cultured in the presence of the ligand. If a significant cell proliferation is detected, after comparison with a negative control in the absence of the ligand, if necessary, the test cDNA is judged to be a "positive clone." Other conditions for culturing transformants can be appropriately selected by one skilled in the art depending on the types of cells into which the vector is inserted and the nature of the fusion protein to be expressed.

[0015] The present invention also relates to a method for isolating a cDNA encoding a peptide that contains the secretory ability. In the method, a cDNA library is ligated into the vector instead of the above test cDNA that is used for detecting cDNA encoding a peptide containing the secretory ability. In one specific embodiment of the invention, cDNAs prepared by using a random primer are ligated with the BstXI adapter and inserted between the two BstXI sites; one is of the vector and the other is inserted into the extracellular cleavage site of the active mpl. The source of the cDNA library is not limited to any specific one, but can be a cell or tissue from which a desired peptide containing the secretory ability is to be isolated. Many standard methods can be used to construct a cDNA library. In the present method, cells judged to be capable of proliferation are selected from the cDNA library-introduced cells. The cDNAs contained in the selected cells are supposed to encode a peptide having the secretory ability. cDNA can be isolated from the cells whose proliferation has been detected by, for example, extracting the genomic DNA or RNA, amplifying the cDNA of interest by PCR using primers designed to encompass the cloning sites (in case of RNA, after converting it into DNA using reverse transcriptase), and recovering the products.

[0016] Whether the recovered cDNA is full length or a fragment, or whether it is a cDNA encoding a novel secretory peptide, can be analyzed by comparing the cDNA sequence with those of the known proteins in the database. If the cDNA is not full length, it is used to screen a secondary cDNA library to isolate a full-length cDNA. The secondary cDNA library can be constructed by a method known to one skilled in the art, such as those described in the literature (Molecular Cloning, A Laboratory Manual, 2nd edition. Sambrook J. et al., (1989) Cold Spring Harbor Laboratory Press, New York).

[0017] A cDNA encoding a peptide with the secretory ability isolated by the method of the present invention can be utilized to produce a recombinant protein that is useful as a medicine or in gene therapy of related diseases. A recombinant protein from the isolated cDNA can be produced by known methods in the art. For example, the cDNA is inserted into an appropriate vector such as pED (Kaufman et al. Nucleic Acids Res. 19:4484-4490 (1991)), pEF-BOS (Mizushima et al. Nucleic Acids Res. 18:5322 (1990)), pXM, pJ13, and pJL4 (Gough et al. EMBO J. 4:645-653 (1985)), the vector is introduced into a host cell, the resulting transformant is cultured to allow it to express a recombinant protein, and the recombinant protein is purified.

Brief Description of the Drawings

[0018]

Figure 1 schematically shows the peptides used for detecting secretory ability and the result of detecting cytokine-independent proliferation ability of BAF/03 cells through the expression of the peptides.

Best Mode for Carrying out the Invention

[0019] The present invention is illustrated in detail below with reference to examples, but is not to be construed as being limited thereto.

Example 1. Vector construction

[0020] In mouse myeloproliferative leukemia virus, env is ligated to the mouse mpl comprising the extracellular domain consisting of 56 amino acids from the transmembrane domain toward the N-terminus, the transmembrane domain, and the intracellular domain. PCR was performed to obtain a cDNA encoding the corresponding region of the human mpl, which is from Leu (449) to the stop codon (636), having the NotI site immediately before the Leu(449) (a single nucleotide insertion) and the SalI site immediately after the stop codon so as to be in the frame of the GM-CSF cDNA shown below. The pBabeX MPL^{M*} (Blood 88:1399-1406. (1996)), in which active human mpl cDNA is cloned, was used as a template. Primers used are listed in Table 1.

Table 1

Not v-mpl (SEQ ID NO: 1)	
	(TGCGGCCGCCCTGGAGCTGCGCCCGCGATCCTGCTACCGTTTA)
	NotI the sequence of mpl
MPL Sal (SEQ ID NO: 2)	
	(GTATGTCGACTCAAGGCTGCTGCCAATAG)
	SalI

[0021] PCR was performed in a reaction mixture containing 10 μg/ml template DNA, 1 μM each primer, 50 U/ml KOD DNA polymerase (TOYOBO), 1 mM $MgCl_2$, 0.2 mM dNTPs, 120 mM Tris-HCl (pH 8), 10 mM KCl, 6 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, and 10 μg/ml BSA by using the GeneAmpPCR System (Perkin Elmer) under the following conditions: denaturation at 98°C for 60 sec, followed by 25 cycles of 98°C for 15 sec, 60°C for 10 sec, and 74°C for 30 sec. The PCR products were analyzed by electrophoresis on an agarose gel, and a gel piece containing a 0.6 kb fragment of interest was excised to extract DNA. The DNA was then phosphorylated at its 5'-termini with T4 polynucleotide kinase (TOYOBO), and ligated by using T4 DNA ligase (TOYOBO) with the pBluescript SK(-) vector (Stratagene) that was pretreated with SmaI (TaKaRa Shuzo) and Bacterial Alkaline Phosphatase (BAP; TaKaRa Shuzo). The nucleotide sequence of the active mpl cDNA inserted in the resulting plasmid was verified with the ABI PRISM 310 Genetic Analyzer (Perkin Elmer). The plasmid was digested with NotI (TaKaRa Shuzo) and SalI (TaKaRa Shuzo), and separated by electrophoresis on an agarose gel to isolate a 0.6 kb fragment. The fragment was ligated with the pMX (Proc. Natl. Acad. Sci. USA 92:9146-9150. (1995)), which was also digested with NotI and SalI, treated with BAP, and purified by agarose gel electrophoresis, using T4 DNA ligase to obtain pMX v-mpl^M. The plasmid pMX v-mpl^M contains a cDNA encoding an active mpl lacking the secretory ability. The nucleotide sequence of the cDNA insert and the amino acid sequence of the peptide encoded by the cDNA are shown in SEQ ID NO: 3 and NO: 4, respectively.

[0022] Next, to obtain a human GM-CSF cDNA in which the stop codon is replaced with a NotI site, PCR was performed by using the pcDSRα 298 hGM-CSF (Proc. Natl. Acad. Sci. USA 82:4360-4364 (1985)) containing the human GM-CSF cDNA as a template. Primers used are shown in Table 2.

Table 2

EcoGMss (SEQ ID NO: 5)	(CGAATTCAAAGTTCTCTGGAGGATG) EcoRI
EcoGM (SEQ ID NO: 6)	(CGAATTCGCCGCCACCATGGCACCCGCCCGCTCGCCC) EcoRI
GM Not (SEQ ID NO: 7)	(AGCGGCCGCCTCCTGGACTGGCTCCCA) NotI

[0023] EcoGM was designed to have the translation initiation codon ATG in place of the Ser (17), and, as in EcoGMss and EcoGM, the EcoRI site and the Kozak consensus sequence (J. Cell Biol. 108:29. (1989)) immediately before the ATG codon. Primer pairs of EcoGMss and GM Not were used in PCR to amplify GM-CSF containing the signal sequence, and EcoGM and GM Not were used to amplify GM-CSF lacking the signal sequence. PCR was performed as described above except for using 55°C for the annealing temperature, and the products were cloned into the pBluescript SK(-). The nucleotide sequence of the DNA inserts was verified by using the ABI PRISM 310 Genetic Analyzer (Perkin Elmer). The plasmids were then digested with EcoRI (TaKaRa) and NotI and inserted into the EcoRI-NotI site of the pMX v-mpl^{M} as described above, and "pMX GM(+)v-mpl^{M}" and "pMX GM(-)v-mpl^{M}" were obtained. The "pMX GM(+)v-mpl^{M}" and "pMX GM(-)v-mpl^{M}" encode a fusion protein between the C-terminal part of the active mpl starting from Leu (449) and the entire GM-CSF with or without the signal sequence, respectively. The nucleotide sequences of their cDNA inserts are shown as SEQ ID NO: 8 and NO: 10, and the amino acid sequences of the proteins encoded by the cDNAs are shown as SEQ ID NO: 9 and NO: 11.

Example 2. Viral infection

[0024] Each of the above plasmids was introduced into packaging cell-line BOSC23 (Proc. Natl. Acad. Sci. USA 90:8392-8396. (1993)) using LipofectAMINE (Life Technologies). BOSC23 cells were plated into 6-cm dishes (CORNING) with Dulbecco's modified Eagle medium (DMEM; Nissui Pharmaceutical) containing 10% fetal calf serum (FCS; JRH Biosciences). After 6-hr incubation, the cells were washed with OPTI-MEM I reduced serum medium (Life Technologies). Separately, LipofectAMINE (18 μl) diluted in 200 μl OPTI-MEM I was mixed with 3 μg samples of each plasmid diluted in 200 μl OPTI-MEM I. The resulting mixtures were allowed to stand at room temperature for 15 minutes, mixed with 1.6 ml OPTI-MEM I, then added to the cells. After 5 hr, 2 ml of DMEM containing 20% FCS was added to the cells, which were incubated for an additional 19 hr. The medium was then replaced with 3 ml of DMEM containing 10% FCS, and the culture supernatant was recovered after 24 hr. Mouse interleukin-3 (IL-3) and 10 μg/ml polybrene (hexadimethrine bromide, Sigma) were added to the culture supernatant containing the recombinant virus, and Ba/F3 cells were suspended therein for infection. After 24 hr of infection, the cells were washed twice in RPMI1640 (Nissui Pharmaceutical) containing 10% FCS lacking mouse IL-3, and the culture was continued in the same medium.

[0025] The cells containing the fusion protein between the entire GM-CSF containing the signal sequence and the active mpl (derived from the pMX GM(+)v-mpl^{M}) grew in the absence of IL-3 as well as those containing the active mpl with the secretory ability. In contrast, the cells containing the fusion protein between the GM-CSF lacking the signal sequence and the active mpl (pMX GM(-)v-mpl^{M}) did not grow as well as control Ba/F3 cells into which no fusion protein expression vector was introduced (Figure 1).

Example 3. Screening

[0026] The following oligonucleotides (Table 3) were synthesized, and their 5'-termini were phosphorylated using T4 polynucleotide kinase. The oligonucleotides were mixed and denatured at 95°C, and then annealed by gradually cooling them to 40°C to prepare the cassette DNA.

Table 3

5'-GGCCCCAGCACAGTGGC-3'	(SEQ ID NO: 12)
5'-GGTCGTGTCACCGCCGG-3'	(SEQ ID NO: 13)

[0027] The pMX GM(-)v-mpl^{M}, which was digested with NotI (TaKaRa) and treated with BAP, was mixed with the cassette and ligated using T4 DNA ligase. The direction of the cassette in the resulting plasmid was verified by DNA sequencing to be in the order of BstXI and NotI (pMX GM(-)v-mpl^{M2}). Total RNA was prepared from the rat neuroblastic cell line MNS70 using the TRIZOL reagent (GIBCO BRL) and passed through the oligo dT column (Pharmacia) to prepare polyA(+) RNA. Double-stranded cDNA was synthesized with the random hexamer contained in the Superscript Choice System (GIBCO BRL). The cDNA was blunt-ended, ligated with the BstXI adapter (Invitrogen), and fractionated by using the SizeSep 400 Spun Column (Pharmacia). The cDNA was then mixed and ligated with the pMX GM(-)v-mpl^{M2} which was digested with BstXI (TaKaRa) and treated with BAP, using T4 DNA ligase. The DNA was introduced into DH10B E. coli (GIBCO BRL) by electroporation using Gene Pulser (BioRad) to construct a cDNA library.

[0028] Plasmids were extracted from the recombinant E. coli containing a cDNA library and purified by using the JETstar column (GENOMED). The library plasmids were introduced into BOSC23 packaging cells by using LipofectAMINE as described above. Mouse IL-3 (10 ng/ml) and 10 μg/ml polybrene (Hexadimethrine Bromide, Sigma) were added to the culture supernatant containing the recombinant virus, and Ba/F3 cells were suspended therein for infection. After 24-hr infection, the cells were washed twice with phosphate buffer and cultured further in RPMI1640 containing 10% FCS. The genomic DNA was prepared from the clones that grew in the absence of IL-3, and PCR was performed using primers designed to encompass the cDNA insertion site to recover the cDNA fragments.

Table 4

5'-GGGGGTGGACCATCCTCTA-3'	(SEQ ID NO: 14)
5'-CGCGCAGCTGTAAACGGTAG-3'	(SEQ ID NO: 15)

[0029] PCR was performed in 50 μl of the reaction mixture containing 500 ng genomic DNA, 500 pM each primer, 2.5 U TaKaRa LA Taq (TaKaRa), 2.5 mM $MgCl_2$, 0.3 mM dNTPs, and the accompanying buffer using the GeneAmpPCR System2400 in the following process: denaturation at 98°C for 60 sec, followed by 30 cycles of 98°C for 20 sec and 68°C for 120 sec. PCR products were separated by electrophoresis on an agarose gel, the gel pieces containing the amplified fragments were excised, and DNA was purified. The nucleotide sequence of the DNA fragments purified from the resulting 190 clones was determined, and 150 clones were found to be cDNAs encoding a known secrete protein or a membrane protein, or its part. The other 40 clones were found to encode unknown secrete proteins. Some of the thus-obtained known secretory proteins are shown in Table 5, where "length" indicates the length of the ORF of the obtained cDNA fragment by the number of amino acid residues. The average length of the clones encoding a known secrete protein was 273 amino acid residues. "Accession number" indicates the accession number in the GenBank protein database. It should be noted that the background in the present method such as detecting a cDNA encoding a protein other than a secrete protein or cDNA that was inserted in the opposite direction was 1% or less.

Table 5

Length	Accession No.	Name
221	1805299	Amyloid precursor
288	416630	Amyloid-like protein 1
350	468563	Amyloid precursor-like protein 2
561	112929	Amyloid A4 protein homologue precursor
161	2494287	o-acetyl GD3 ganglioside synthase
176	2507439	Syndecan 3 (heparan sulfate proteoglycan core protein)
218	118115	Cyr61 protein (growth factor binding protein)
382	3219172	collagen alpha 1(V)

Table 5 (continued)

Length	Accession No.	Name
286	461671	collagen alpha 1 type 1
159	1082724	Prostacyclin-stimulating factor
259	1777354	SHPS-1, BIT
254	205167	120 kDa sialoglycoprotein (a hepatic lysosomal membrane protein)
482	1708023	K-glypican
224	1139548	Seizure-related gene product 6 type2 precursor
105	135818	G-protein coupled thrombin receptor
482	129731	Protein Disulfide Isomerase
322	1172451	perlecan (basement membrane heparin sulfate proteoglycan)
165	1709256	neurocan (proteoglycan core protein precursor)
264	2367641	neuropilin-2 (semaphorin III receptor)
211	126638	Lysyl oxidase
308	2627143	Neural cadherin
459	3123675	Notch
140	1718156	Vascular endothelial growth factor
534	627989	Endothelin-converting enzyme
89	114393	Sodium/potassium-transporting ATPase beta-1 chain

Industrial Applicability

[0030] The present invention provides a method for detecting and isolating a cDNA encoding a secretory peptide using a peptide capable of triggering cell proliferation through its dimerization on the cell surface but lacking the secretory ability. Since the method utilizes cell proliferation as an index for detection, it is extremely easy and sensitive. Moreover, compared to the conventional methods that enable detecting a short DNA fragment, this method enables detecting and isolating a cDNA containing a longer peptide coding region, and thus provides more information from the first isolated clones. In addition, the method enables detecting and isolating secretory proteins including type I and type II membrane proteins.

SEQUENCE LISTING

<110> CHUGAI SEIYAKU KABUSHIKI KAISHA
KITAMURA, Toshio

<120> SIGNAL SEQUENCE TRAPPING METHOD

<130> C1-902PCT

<150> JP 1997-324912
<151> 1997-11-26

<160> 15

<210> 1
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificially synthesized primer sequence

<400> 1
tgcggccgcc ctggagctgc gcccgcgatc ctgctaccgt tta 43

<210> 2
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificially synthesized primer sequence

<400> 2

gtatgtcgac tcaaggctgc tgccaatag 29

<210> 3
<211> 579
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificially manipulated human mpl sequence

<220>
<221> CDS
<222> (1)..(573)

<400> 3

```
gcg gcc gcc ctg gag ctg cgc ccg cga tct cgc tac cgt tta cag ctg      48
Ala Ala Ala Leu Glu Leu Arg Pro Arg Ser Arg Tyr Arg Leu Gln Leu
1               5                  10                  15

cgc gcc agg ctc aac ggc ccc acc tac caa ggt ccc tgg agc tcg tgg      96
Arg Ala Arg Leu Asn Gly Pro Thr Tyr Gln Gly Pro Trp Ser Ser Trp
            20                  25                  30

tcg gac cca act agg gtg gag acc gcc acc gag acc gcc tgg atc tcc     144
Ser Asp Pro Thr Arg Val Glu Thr Ala Thr Glu Thr Ala Trp Ile Ser
        35                  40                  45

ttg gtg acc gct ctg cat cta gtg ctg ggc ctc aac gcc gtc ctg ggc     192
Leu Val Thr Ala Leu His Leu Val Leu Gly Leu Asn Ala Val Leu Gly
    50                  55                  60

ctg ctg ctg ctg agg tgg cag ttt cct gca cac tac agg aga ctg agg     240
Leu Leu Leu Leu Arg Trp Gln Phe Pro Ala His Tyr Arg Arg Leu Arg
65                  70                  75                  80
```

cat gcc ctg tgg ccc tca ctt cca gac ctg cac cgg gtc cta ggc cag 288
His Ala Leu Trp Pro Ser Leu Pro Asp Leu His Arg Val Leu Gly Gln
85 90 95

tac ctt agg gac act gca gcc ctg agc ccg ccc aag gcc aca gtc tca 336
Tyr Leu Arg Asp Thr Ala Ala Leu Ser Pro Pro Lys Ala Thr Val Ser
100 105 110

gat acc tgt gaa gaa gtg gaa ccc agc ctc ctt gaa atc ctc ccc aag 384
Asp Thr Cys Glu Glu Val Glu Pro Ser Leu Leu Glu Ile Leu Pro Lys
115 120 125

tcc tca gag agg act cct ttg ccc ctg tgt tcc tcc cag gcc cag atg 432
Ser Ser Glu Arg Thr Pro Leu Pro Leu Cys Ser Ser Gln Ala Gln Met
130 135 140

gac tac cga aga ttg cag cct tct tgc ctg ggg acc atg ccc ctg tct 480
Asp Tyr Arg Arg Leu Gln Pro Ser Cys Leu Gly Thr Met Pro Leu Ser
145 150 155 160

gtg tgc cca ccc atg gct gag tca ggg tcc tgc tgt acc acc cac att 528
Val Cys Pro Pro Met Ala Glu Ser Gly Ser Cys Cys Thr Thr His Ile
165 170 175

gcc aac cat tcc tac cta cca cta agc tat tgg cag cag cct 570
Ala Asn His Ser Tyr Leu Pro Leu Ser Tyr Trp Gln Gln Pro
180 185 190

tgagtcgac 579

<210> 4
<211> 190
<212> PRT
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificially manipulated human mpl sequence

<400> 4

```
Ala Ala Ala Leu Glu Leu Arg Pro Arg Ser Arg Tyr Arg Leu Gln Leu
1               5                  10                  15

Arg Ala Arg Leu Asn Gly Pro Thr Tyr Gln Gly Pro Trp Ser Ser Trp
            20                  25                  30

Ser Asp Pro Thr Arg Val Glu Thr Ala Thr Glu Thr Ala Trp Ile Ser
        35                  40                  45

Leu Val Thr Ala Leu His Leu Val Leu Gly Leu Asn Ala Val Leu Gly
    50                  55                  60

Leu Leu Leu Leu Arg Trp Gln Phe Pro Ala His Tyr Arg Arg Leu Arg
65                  70                  75                  80

His Ala Leu Trp Pro Ser Leu Pro Asp Leu His Arg Val Leu Gly Gln
                85                  90                  95

Tyr Leu Arg Asp Thr Ala Ala Leu Ser Pro Pro Lys Ala Thr Val Ser
            100                 105                 110

Asp Thr Cys Glu Glu Val Glu Pro Ser Leu Leu Glu Ile Leu Pro Lys
        115                 120                 125

Ser Ser Glu Arg Thr Pro Leu Pro Leu Cys Ser Ser Gln Ala Gln Met
    130                 135                 140

Asp Tyr Arg Arg Leu Gln Pro Ser Cys Leu Gly Thr Met Pro Leu Ser
145                 150                 155                 160
```

Val Cys Pro Pro Met Ala Glu Ser Gly Ser Cys Cys Thr Thr His Ile
165 170 175

Ala Asn His Ser Tyr Leu Pro Leu Ser Tyr Trp Gln Gln Pro
180 185 190

<210> 5
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificially synthesized primer sequence

<400> 5
cgaattcaaa gttctctgga ggatg 25

<210> 6
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificially synthesized primer sequence

<400> 6
cgaattcgcc gccaccatgg cacccgcccg ctcgccc 37

<210> 7
<211> 27
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificially synthesized primer sequence

<400> 7

agcggccgcc tcctggactg gctccca 27

<210> 8

<211> 1032

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificially synthesized human GM-CSF-human mpl fusion gene sequence

<220>

<221> sig_peptide

<222> (22)..(72)

<220>

<221> CDS

<222> (22)..(1026)

<400> 8

gaattcaaag ttctctggag g atg tgg ctg cag agc ctg ctg ctc ttg ggc 51
Met Trp Leu Gln Ser Leu Leu Leu Leu Gly
1 5 10

act gtg gcc tgc agc atc tct gca ccc gcc cgc tcg ccc agc ccc agc 99
Thr Val Ala Cys Ser Ile Ser Ala Pro Ala Arg Ser Pro Ser Pro Ser
15 20 25

acg cag ccc tgg gag cat gtg aat gcc atc cag gag gcc cgg cgt ctc 147
Thr Gln Pro Trp Glu His Val Asn Ala Ile Gln Glu Ala Arg Arg Leu
30 35 40

```
ctg aac ctg agt aga gac act gct gct gag atg aat gaa aca gta gaa      195
Leu Asn Leu Ser Arg Asp Thr Ala Ala Glu Met Asn Glu Thr Val Glu
        45                  50                  55

gtc atc tca gaa atg ttt gac ctc cag gag ccg acc tgc cta cag acc      243
Val Ile Ser Glu Met Phe Asp Leu Gln Glu Pro Thr Cys Leu Gln Thr
    60                  65                  70

cgc ctg gag ctg tac aag cag ggc ctg cgg ggc agc ctc acc aag ctc      291
Arg Leu Glu Leu Tyr Lys Gln Gly Leu Arg Gly Ser Leu Thr Lys Leu
75                  80                  85                  90

aag ggc ccc ttg acc atg atg gcc agc cac tac aag cag cac tgc cct      339
Lys Gly Pro Leu Thr Met Met Ala Ser His Tyr Lys Gln His Cys Pro
                95                  100                 105

cca acc ccg gaa act tcc tgt gca acc cag att atc acc ttt gaa agt      387
Pro Thr Pro Glu Thr Ser Cys Ala Thr Gln Ile Ile Thr Phe Glu Ser
            110                 115                 120

ttc aaa gag aac ctg aag gac ttt ctg ctt gtc atc ccc ttt gac tgc      435
Phe Lys Glu Asn Leu Lys Asp Phe Leu Leu Val Ile Pro Phe Asp Cys
        125                 130                 135

tgg gag cca gtc cag gag gcg gcc gcc ctg gag ctg cgc ccg cga tct      483
Trp Glu Pro Val Gln Glu Ala Ala Ala Leu Glu Leu Arg Pro Arg Ser
    140                 145                 150

cgc tac cgt tta cag ctg cgc gcc agg ctc aac ggc ccc acc tac caa      531
Arg Tyr Arg Leu Gln Leu Arg Ala Arg Leu Asn Gly Pro Thr Tyr Gln
155                 160                 165                 170

ggt ccc tgg agc tcg tgg tcg gac cca act agg gtg gag acc gcc acc      579
Gly Pro Trp Ser Ser Trp Ser Asp Pro Thr Arg Val Glu Thr Ala Thr
                175                 180                 185
```

```
gag acc gcc tgg atc tcc ttg gtg acc gct ctg cat cta gtg ctg ggc      627
Glu Thr Ala Trp Ile Ser Leu Val Thr Ala Leu His Leu Val Leu Gly
            190                 195                 200

ctc aac gcc gtc ctg ggc ctg ctg ctg ctg agg tgg cag ttt cct gca      675
Leu Asn Ala Val Leu Gly Leu Leu Leu Leu Arg Trp Gln Phe Pro Ala
        205                 210                 215

cac tac agg aga ctg agg cat gcc ctg tgg ccc tca ctt cca gac ctg      723
His Tyr Arg Arg Leu Arg His Ala Leu Trp Pro Ser Leu Pro Asp Leu
    220                 225                 230

cac cgg gtc cta ggc cag tac ctt agg gac act gca gcc ctg agc ccg      771
His Arg Val Leu Gly Gln Tyr Leu Arg Asp Thr Ala Ala Leu Ser Pro
235                 240                 245                 250

ccc aag gcc aca gtc tca gat acc tgt gaa gaa gtg gaa ccc agc ctc      819
Pro Lys Ala Thr Val Ser Asp Thr Cys Glu Glu Val Glu Pro Ser Leu
                255                 260                 265

ctt gaa atc ctc ccc aag tcc tca gag agg act cct ttg ccc ctg tgt      867
Leu Glu Ile Leu Pro Lys Ser Ser Glu Arg Thr Pro Leu Pro Leu Cys
            270                 275                 280

tcc tcc cag gcc cag atg gac tac cga aga ttg cag cct tct tgc ctg      915
Ser Ser Gln Ala Gln Met Asp Tyr Arg Arg Leu Gln Pro Ser Cys Leu
        285                 290                 295

ggg acc atg ccc ctg tct gtg tgc cca ccc atg gct gag tca ggg tcc      963
Gly Thr Met Pro Leu Ser Val Cys Pro Pro Met Ala Glu Ser Gly Ser
    300                 305                 310

tgc tgt acc acc cac att gcc aac cat tcc tac cta cca cta agc tat     1011
Cys Cys Thr Thr His Ile Ala Asn His Ser Tyr Leu Pro Leu Ser Tyr
315                 320                 325                 330
```

tgg cag cag cct tgagtcgac 1032
Trp Gln Gln Pro

<210> 9
<211> 334
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificially synthesized human GM-CSF-human mpl fusion protein sequence

<400> 9

Met Trp Leu Gln Ser Leu Leu Leu Leu Gly
1 5 10

Thr Val Ala Cys Ser Ile Ser Ala Pro Ala Arg Ser Pro Ser Pro Ser
15 20 25

Thr Gln Pro Trp Glu His Val Asn Ala Ile Gln Glu Ala Arg Arg Leu
30 35 40

Leu Asn Leu Ser Arg Asp Thr Ala Ala Glu Met Asn Glu Thr Val Glu
45 50 55

Val Ile Ser Glu Met Phe Asp Leu Gln Glu Pro Thr Cys Leu Gln Thr
60 65 70

Arg Leu Glu Leu Tyr Lys Gln Gly Leu Arg Gly Ser Leu Thr Lys Leu
75 80 85 90

Lys Gly Pro Leu Thr Met Met Ala Ser His Tyr Lys Gln His Cys Pro
95 100 105

```
Pro Thr Pro Glu Thr Ser Cys Ala Thr Gln Ile Ile Thr Phe Glu Ser
            110                 115                 120

Phe Lys Glu Asn Leu Lys Asp Phe Leu Leu Val Ile Pro Phe Asp Cys
        125                 130                 135

Trp Glu Pro Val Gln Glu Ala Ala Ala Leu Glu Leu Arg Pro Arg Ser
    140                 145                 150

Arg Tyr Arg Leu Gln Leu Arg Ala Arg Leu Asn Gly Pro Thr Tyr Gln
155                 160                 165                 170

Gly Pro Trp Ser Ser Trp Ser Asp Pro Thr Arg Val Glu Thr Ala Thr
                175                 180                 185

Glu Thr Ala Trp Ile Ser Leu Val Thr Ala Leu His Leu Val Leu Gly
            190                 195                 200

Leu Asn Ala Val Leu Gly Leu Leu Leu Leu Arg Trp Gln Phe Pro Ala
        205                 210                 215

His Tyr Arg Arg Leu Arg His Ala Leu Trp Pro Ser Leu Pro Asp Leu
    220                 225                 230

His Arg Val Leu Gly Gln Tyr Leu Arg Asp Thr Ala Ala Leu Ser Pro
235                 240                 245                 250

Pro Lys Ala Thr Val Ser Asp Thr Cys Glu Glu Val Glu Pro Ser Leu
                255                 260                 265

Leu Glu Ile Leu Pro Lys Ser Ser Glu Arg Thr Pro Leu Pro Leu Cys
            270                 275                 280

Ser Ser Gln Ala Gln Met Asp Tyr Arg Arg Leu Gln Pro Ser Cys Leu
        285                 290                 295
```

Gly Thr Met Pro Leu Ser Val Cys Pro Pro Met Ala Glu Ser Gly Ser
300 305 310

Cys Cys Thr Thr His Ile Ala Asn His Ser Tyr Leu Pro Leu Ser Tyr
315 320 325 330

Trp Gln Gln Pro

<210> 10
<211> 1032
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificially synthesized human GM-CSF-human mpl fusion gene sequence

<220>
<221> CDS
<222> (16)..(972)

<400> 10
gaattcaaag ttctctggag g atg tgg ctg cag agc ctg ctg ctc ttg ggc 51
Met Trp Leu Gln Ser Leu Leu Leu Leu Gly
1 5 10

act gtg gcc tgc agc atc tct gca ccc gcc cgc tcg ccc agc ccc agc 99
Thr Val Ala Cys Ser Ile Ser Ala Pro Ala Arg Ser Pro Ser Pro Ser
15 20 25

acg cag ccc tgg gag cat gtg aat gcc atc cag gag gcc cgg cgt ctc 147
Thr Gln Pro Trp Glu His Val Asn Ala Ile Gln Glu Ala Arg Arg Leu
30 35 40

```
ctg aac ctg agt aga gac act gct gct gag atg aat gaa aca gta gaa      195
Leu Asn Leu Ser Arg Asp Thr Ala Ala Glu Met Asn Glu Thr Val Glu
        45                  50                  55

gtc atc tca gaa atg ttt gac ctc cag gag ccg acc tgc cta cag acc      243
Val Ile Ser Glu Met Phe Asp Leu Gln Glu Pro Thr Cys Leu Gln Thr
    60                  65                  70

cgc ctg gag ctg tac aag cag ggc ctg cgg ggc agc ctc acc aag ctc      291
Arg Leu Glu Leu Tyr Lys Gln Gly Leu Arg Gly Ser Leu Thr Lys Leu
75                  80                  85                  90

aag ggc ccc ttg acc atg atg gcc agc cac tac aag cag cac tgc cct      339
Lys Gly Pro Leu Thr Met Met Ala Ser His Tyr Lys Gln His Cys Pro
                95                  100                 105

cca acc ccg gaa act tcc tgt gca acc cag att atc acc ttt gaa agt      387
Pro Thr Pro Glu Thr Ser Cys Ala Thr Gln Ile Ile Thr Phe Glu Ser
            110                 115                 120

ttc aaa gag aac ctg aag gac ttt ctg ctt gtc atc ccc ttt gac tgc      435
Phe Lys Glu Asn Leu Lys Asp Phe Leu Leu Val Ile Pro Phe Asp Cys
        125                 130                 135

tgg gag cca gtc cag gag gcg gcc gcc ctg gag ctg cgc ccg cga tct      483
Trp Glu Pro Val Gln Glu Ala Ala Ala Leu Glu Leu Arg Pro Arg Ser
    140                 145                 150

cgc tac cgt tta cag ctg cgc gcc agg ctc aac ggc ccc acc tac caa      531
Arg Tyr Arg Leu Gln Leu Arg Ala Arg Leu Asn Gly Pro Thr Tyr Gln
155                 160                 165                 170

ggt ccc tgg agc tcg tgg tcg gac cca act agg gtg gag acc gcc acc      579
Gly Pro Trp Ser Ser Trp Ser Asp Pro Thr Arg Val Glu Thr Ala Thr
                175                 180                 185
```

```
gag acc gcc tgg atc tcc ttg gtg acc gct ctg cat cta gtg ctg ggc      627
Glu Thr Ala Trp Ile Ser Leu Val Thr Ala Leu His Leu Val Leu Gly
            190                 195                 200

ctc aac gcc gtc ctg ggc ctg ctg ctg ctg agg tgg cag ttt cct gca      675
Leu Asn Ala Val Leu Gly Leu Leu Leu Leu Arg Trp Gln Phe Pro Ala
        205                 210                 215

cac tac agg aga ctg agg cat gcc ctg tgg ccc tca ctt cca gac ctg      723
His Tyr Arg Arg Leu Arg His Ala Leu Trp Pro Ser Leu Pro Asp Leu
    220                 225                 230

cac cgg gtc cta ggc cag tac ctt agg gac act gca gcc ctg agc ccg      771
His Arg Val Leu Gly Gln Tyr Leu Arg Asp Thr Ala Ala Leu Ser Pro
235                 240                 245                 250

ccc aag gcc aca gtc tca gat acc tgt gaa gaa gtg gaa ccc agc ctc      819
Pro Lys Ala Thr Val Ser Asp Thr Cys Glu Glu Val Glu Pro Ser Leu
                255                 260                 265

ctt gaa atc ctc ccc aag tcc tca gag agg act cct ttg ccc ctg tgt      867
Leu Glu Ile Leu Pro Lys Ser Ser Glu Arg Thr Pro Leu Pro Leu Cys
            270                 275                 280

tcc tcc cag gcc cag atg gac tac cga aga ttg cag cct tct tgc ctg      915
Ser Ser Gln Ala Gln Met Asp Tyr Arg Arg Leu Gln Pro Ser Cys Leu
        285                 290                 295

ggg acc atg ccc ctg tct gtg tgc cca ccc atg gct gag tca ggg tcc      963
Gly Thr Met Pro Leu Ser Val Cys Pro Pro Met Ala Glu Ser Gly Ser
    300                 305                 310

tgc tgt acc acc cac att gcc aac cat tcc tac cta cca cta agc tat     1011
Cys Cys Thr Thr His Ile Ala Asn His Ser Tyr Leu Pro Leu Ser Tyr
315                 320                 325                 330
```

tgg cag cag cct tgagtcgac 1032
Trp Gln Gln Pro

<210> 11
<211> 334
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificially synthesized human GM-CSF-human mpl fusion protein sequence

<400> 11

Met Trp Leu Gln Ser Leu Leu Leu Leu Gly
1 5 10

Thr Val Ala Cys Ser Ile Ser Ala Pro Ala Arg Ser Pro Ser Pro Ser
15 20 25

Thr Gln Pro Trp Glu His Val Asn Ala Ile Gln Glu Ala Arg Arg Leu
30 35 40

Leu Asn Leu Ser Arg Asp Thr Ala Ala Glu Met Asn Glu Thr Val Glu
45 50 55

Val Ile Ser Glu Met Phe Asp Leu Gln Glu Pro Thr Cys Leu Gln Thr
60 65 70

Arg Leu Glu Leu Tyr Lys Gln Gly Leu Arg Gly Ser Leu Thr Lys Leu
75 80 85 90

Lys Gly Pro Leu Thr Met Met Ala Ser His Tyr Lys Gln His Cys Pro
95 100 105

```
Pro Thr Pro Glu Thr Ser Cys Ala Thr Gln Ile Ile Thr Phe Glu Ser
                110                 115                 120

Phe Lys Glu Asn Leu Lys Asp Phe Leu Leu Val Ile Pro Phe Asp Cys
        125                 130                 135

Trp Glu Pro Val Gln Glu Ala Ala Ala Leu Glu Leu Arg Pro Arg Ser
    140                 145                 150

Arg Tyr Arg Leu Gln Leu Arg Ala Arg Leu Asn Gly Pro Thr Tyr Gln
155                 160                 165                 170

Gly Pro Trp Ser Ser Trp Ser Asp Pro Thr Arg Val Glu Thr Ala Thr
                175                 180                 185

Glu Thr Ala Trp Ile Ser Leu Val Thr Ala Leu His Leu Val Leu Gly
            190                 195                 200

Leu Asn Ala Val Leu Gly Leu Leu Leu Leu Arg Trp Gln Phe Pro Ala
        205                 210                 215

His Tyr Arg Arg Leu Arg His Ala Leu Trp Pro Ser Leu Pro Asp Leu
    220                 225                 230

His Arg Val Leu Gly Gln Tyr Leu Arg Asp Thr Ala Ala Leu Ser Pro
235                 240                 245                 250

Pro Lys Ala Thr Val Ser Asp Thr Cys Glu Glu Val Glu Pro Ser Leu
                255                 260                 265

Leu Glu Ile Leu Pro Lys Ser Ser Glu Arg Thr Pro Leu Pro Leu Cys
            270                 275                 280

Ser Ser Gln Ala Gln Met Asp Tyr Arg Arg Leu Gln Pro Ser Cys Leu
        285                 290                 295
```

Gly Thr Met Pro Leu Ser Val Cys Pro Pro Met Ala Glu Ser Gly Ser
300 305 310

Cys Cys Thr Thr His Ile Ala Asn His Ser Tyr Leu Pro Leu Ser Tyr
315 320 325 330

Trp Gln Gln Pro

<210> 12
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificially synthesized casette sequence

<400> 12
ggccccagca cagtggc 17

<210> 13
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: artificially synthesized casette sequence

<400> 13
ggccgccact gtgctgg 17

<210> 14
<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificially synthesized primer sequence

<400> 14

ggggtggac catcctcta 19

<210> 15

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: artificially synthesized primer sequence

<400> 15

cgcgcagctg taaacggtag 20

Claims

1. A peptide capable of inducing cell proliferation through dimerization on the cell surface and lacking the secretory ability.
2. The peptide of claim 1, wherein the peptide is derived from a cytokine receptor.
3. The peptide of claim 1, wherein the peptide is derived from mpl.
4. The peptide of claim 2 or 3, wherein the peptide is ligand-independent.
5. The peptide of claim 1, wherein the peptide comprises the amino acid sequence of SEQ ID NO: 4.
6. A DNA encoding the peptide of any one of claims 1 to 5.
7. A vector containing the DNA of claim 6 and a cloning site for cDNA at the 5'-upstream region of the DNA.
8. The vector of claim 7, wherein the vector is derived from a retrovirus.
9. The vector of claims 7 or 8, wherein a cDNA is inserted into the 5'-upstream of the DNA of claim 6.

10. A cell carrying the vector of claim 9.

11. The cell of claim 10, wherein the cell is a mammalian cell.

12. A method for examining whether a peptide encoded by a cDNA to be tested contains the secretory ability, the method comprising

(a) ligating the test cDNA with the vector of claim 7;
(b) introducing the vector prepared in (a) into a cell; and
(c) culturing the transformant prepared in (b), and detecting the cell proliferation ability.

13. A method for isolating a cDNA encoding a peptide with the secretory ability, the method comprising

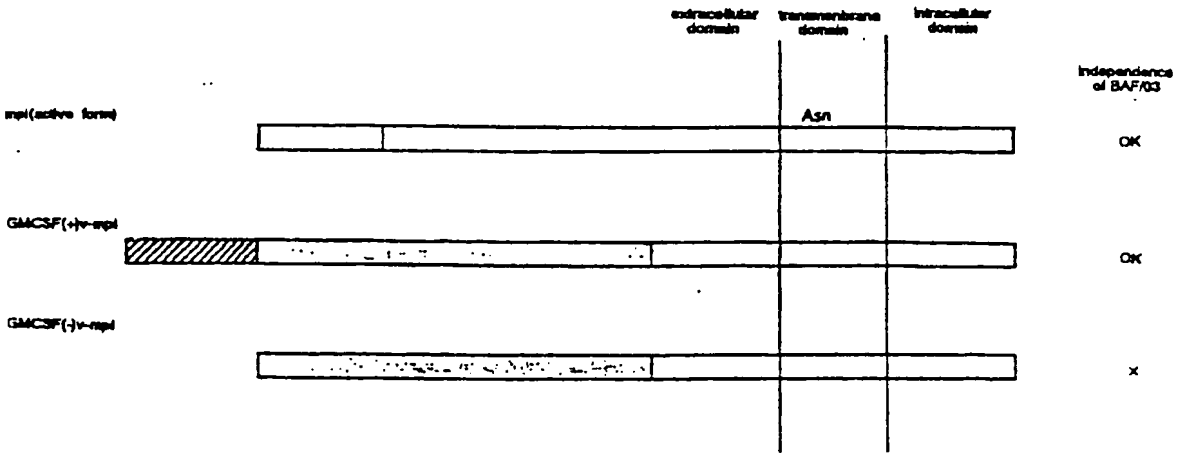
(a) ligating a cDNA library with the vector of claim 7;
(b) introducing the vector prepared in (a) into a cell;
(c) culturing the transformant prepared in (b), and detecting the cell proliferation ability; and
(d) selecting a positive cell that is judged to have cell proliferation ability in (c), and isolating the cDNA from said cell.

14. The method of claims 12 or 13, wherein the vector is derived from a retrovirus and the cell to be introduced with the vector is a mammalian cell.

15. A cDNA encoding a peptide with the secretory ability, which is isolated by the method of claim 13.

16. A peptide encoded by the cDNA of claim 15.

Fig.1


extracellular domain
transmembrane domain
intracellular domain
Independance of BAF/03
mpl(active form)
Asn
OK
GMCSF(+)v-mpl
OK
GMCSF(-)v-mpl
×

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP98/05326

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl[6] C07K14/715, C12N15/12, C12N5/16, C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int.Cl[6] C07K14/715, C12N15/12, C12N5/16, C12Q1/68

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
WPI (DIALOG), BIOSIS (DIALOG)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X/Y	ISABELLE VIGON et al., "Molecular cloning and characterization of MPL, the human homolog of the v-mpl oncogene ···" Proc. Natl. Acad. Sci. USA, 1992, 89, 5640-5644	1-3, 5-11, 15, 16/4
X	KAZUHIRO HAYASHIDA et al., "Molecular cloning of a second subunit of the receptor for human granulocyte-macrophage colony-stimulating factor ···" Proc. Natl. Acad. Sci. USA, 1990, 87, 9655-9659	1, 2, 6-11, 15, 16
Y	MAYUMI ONISHI et al., "Identification of an oncogenic form of the Thrombopoietin receptor MPL ···" Blood, 1996, 88, 1399-1406	4

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 January, 1999 (19. 01. 99)	26 January, 1999 (26. 01. 99)
Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

Form PCT/ISA/210 (second sheet) (July 1992)